Developing Decision Model by Mining Historical Prices Data of Infosys for Stock Market Prediction

Sanjeev Gour

Department of Computer Science, Career College, Bhopal, India

e-mail-sunj129@gmail.com

Available online at: www.ijcseonline.org

Abstract-**Stock market analysis is the process of analyzing and monitoring stocks so it is also a process of calculating the future trends on the basis of historical trends. This whole concept is volatile, as the stock prices having the tendency to rise and fall. However, we know that there is a defined pattern in insight of any sequenced event therefore we can extract some hidden pattern thorough analysis. In this paper we have developed a decision support model to classify and predict the stock market by data mining techniques like classification and prediction. In this way we have developed some decision rules as model to increase the probability of right decision so that an investor can took profit in the stock investment. in this study we analyze the historical price data of the specific industry group Named Infosys Pvt. Ltd. to make sure that the investors is moving with right decision   in order to increase the possibility of profit in their investments. Therefore the main task is to predict and classify the stock prices of Infosys Company on the basis of past prices.**

Keywords: Classification, Data Mining, Prediction, Stock market.

I.   INTRODUCTION

To predict and classify the future stock prices, we used the historical stock prices of particular company listed in the stock market [1]. Stock market prediction is the process of determining the future value of a company stock on an exchange and the successful prediction value of stock price can be beneficial for the investors. The stock prices fully reflect all currently available stock data and this can be a good-market hypothesis [2]. Nowadays there are many technologies and methodologies have been used to analyze the historical stock data of any company and the successful prediction could yield profit to investors [3, 4]. Many prediction methods can be fall into two categories: Basic or fundamental analysis and technical   or technological analysis. Fundamental Analysis emphasis the company that underlies the stock itself. It analyzes company's past performance. Technical analysis seeks to determine the future price of a stock based solely on the past price. One of the most common fields of analyzing this type of data is Data mining.

Data mining Technology: There are numerous data mining techniques to extract useful information from the historical stock price data of particular company [5]. Many successful stock market predictions have done through data mining methods. The most prominent technique involves the use of classification and prediction generic algorithms and neural networks [6, 7]. As the data mining methods have great capability to extract hidden pattern from the large dataset, this is one of the major reason that the current research uses these methodologies to build a decision model and also focuses in

the stock forecasting area to improve the accuracy of stock trading forecast. This paper introduces various data mining techniques and supports the decision-making for stock trades.

II.   ROLE OF DATA MINING IN STOCK MARKET

As we know that the historic data include the essential information for predicting the future direction. Today many technologies which are designed to help investors to discover hidden patterns from the historic data [8]. Data mining is one of technology nowadays used frequently to analyzing big past price data of particular company/exchange for predicting the stocks market Prices [9]. These data mining methods help the investors in the stock market to decide the better timing for buying or selling stocks based on the extracted pattern or information from the historical prices of stocks [10].

Livermore [11] believed that stock trends follow a trend line that can be used to forecast both in the long- and short-term. He published this particular idea in "How to Trade in Stock" in 1940 .using stock data he concluded that stock-group behavior was an important indication to overall market direction, whether they are big or small.

III.   ABOUT DATASET

For the purpose of this study, 6 years INFOSYS LTD INR5 (INFY.NS) monthly stock data employed from NSE - National Stock Exchange of India Ltd. [12].The data employed in this study contain variables open price, high price, low price and close price of Infosys index. The data set encompassed the trading months from 1-1-2010 to 30-12-

2015. Stock market price data is very hard to interpret because prices of stock are very large in numerically and not easy for computation. So the data need to be change in symbolic representation and therefore we used the variables Open Price, High Price, Low Price and Close Price to represent this conversion, which is shown in Table 2 and table 3. The Table 2 shows symbolic conversion of data
Data format:

values. For this, data values are classified under the below average and the above average. The data values below average comes under the category of falling price and the values which are above average comes under the category of rising price. The Table 2 & Table 3 shows this classification for Open/High and Low/Close attribute.

Table 1: Historical Dataset after Preprocessing

Date	ID	OPEN	HIGH	LOW	CLOSE	VOLUME	STATUS1	STATUS2
12/01/15	1	1088	1110	1022.54999	1105.4002	2708400	Fall	Fall
……	…	…..	…..	……..	………..	…….	…….	……
10/01/14	15	3722	4066	3572.69995	4051.2399	5283000	Rise	Rise
…….	…	……	……	…………..	…….	………….	…….	…….
02/01/10	72	2447	2630	2329	2601.948	4146700	Fall	Fall

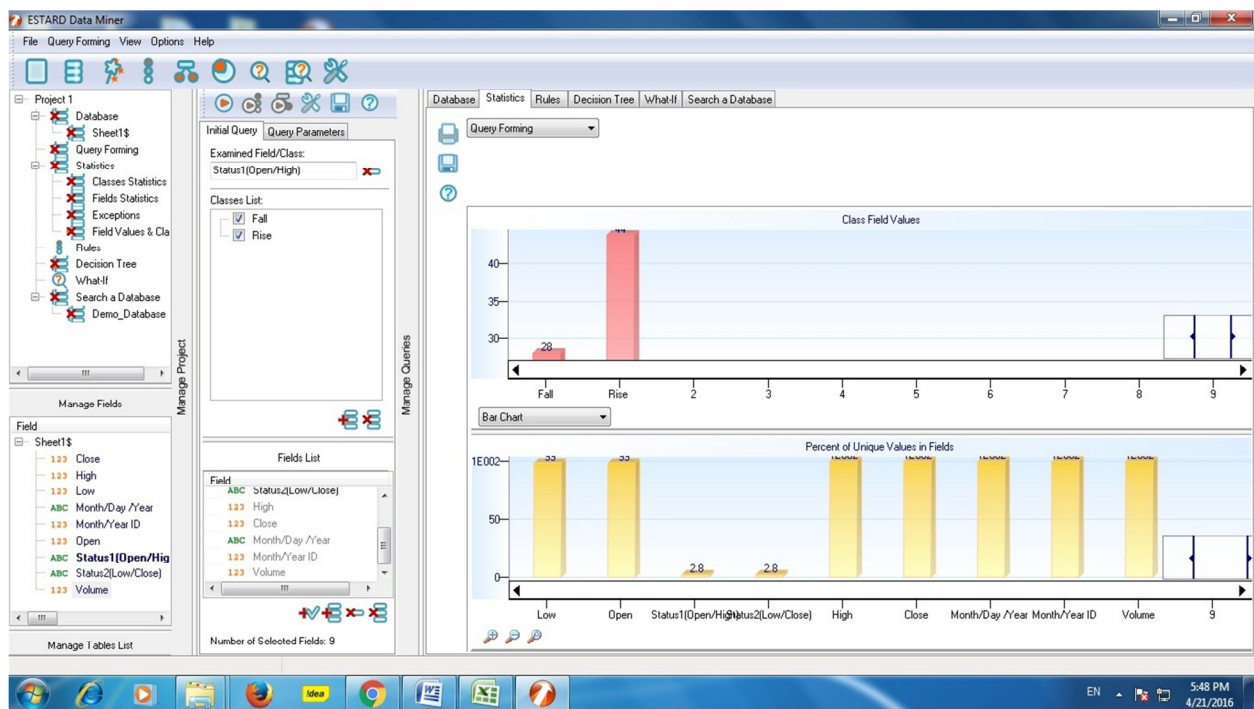Figure1: Attribute view in Estard Data miner.

Table 2: Attribute Status1 (open /high) for category
Fall/Rise

Attribute	OPEN	HIGH
Values	lies b/w 989-4076.88	Values lies b/w 1110-4281.64
Average	2665.411	2823.58
	Open_Fall	High_Fall
	Open_Rise	High_Rise

Table 3: Attribute Status2 (low /close) for category
Fall/Rise

Attribute	LOW	CLOSE
Values	lies b/w 932.65-3572.69	lies b/w 984.35-4051.24

Average	2478.38	2641.811
	Low_Fall	Close_Fall
	Low_Rise	Close_Rise

IV. SOFTWARE TOOL FOR ANALYSIS

ESTARD Data Miner (EDM) is a data mining tool, able to discover most unexpected hidden information in your data. We have seen many databases include data that is assembled for many years. These databases (also called data warehouses) can become a valuable source of new knowledge for our analysis. The latest business intelligence techniques were incorporated into ESTARD Data Miner for generating automated data analysis. User-friendly interface and wizards allow starting working with the tool in a few clicks. We have used ESTARD Version 3.0 Free Available Data mining Software. Rules (also called if-then rules, or production rules) and decision trees are powerful data mining methods allowing analyzing hidden correlations in your data. In ESTARD Data Miner you have a possibility to use both these methods. With the help of these methods you can create models that will describe your data and will help in further decision making. ESTARD Data Miner is a powerful tool that gives you technologies for understanding your business processes, for analyzing and predicting what to expect in future.

V.   EXPERIMENTAL PROCESS

The first step for the Data Mining process is selecting the target database. It can be any database containing information you would like to use for data mining. This database is called the Learning database. After loading the learning database the list of tables and table fields' details will be displayed on the "Database" page. To perform Rules query or Decision Tree query first you have to perform the initial Statistics query. To create rules or trees you can use Query Wizard. Then we select the examined "class" field. In our case we have select attribute Satus1 (open/high) and Status2 (Low/close) and their values have two values "fall" and "Rise". Then we have Select classes and fields to use for rules creation and finally Create Decision rules and tree as a Model. (Fig-2 and Fig- 3). If-then Rules, also called Decision Rules or Production Rules are a basic structure in data mining and expert systems. This method of knowledge representation is simple and easy to use for data analysis. Each rule defines a piece of overall knowledge. Rules can be edited, removed or added independently from each other. Decision trees are an excellent tool in decision-making and data mining systems. They can be of good service to any analyst, manager or scientist. Decision Tree is a data mining method that allows creating easy to understand and use decision models. The decisions are represented as the nodes of the tree model. The results of the decisions are outputted as the final "leafs" of the tree. Now after we have created rules or decision tree also we use the obtained data for WHAT-IF analysis to predict the examined class (Fig-4).

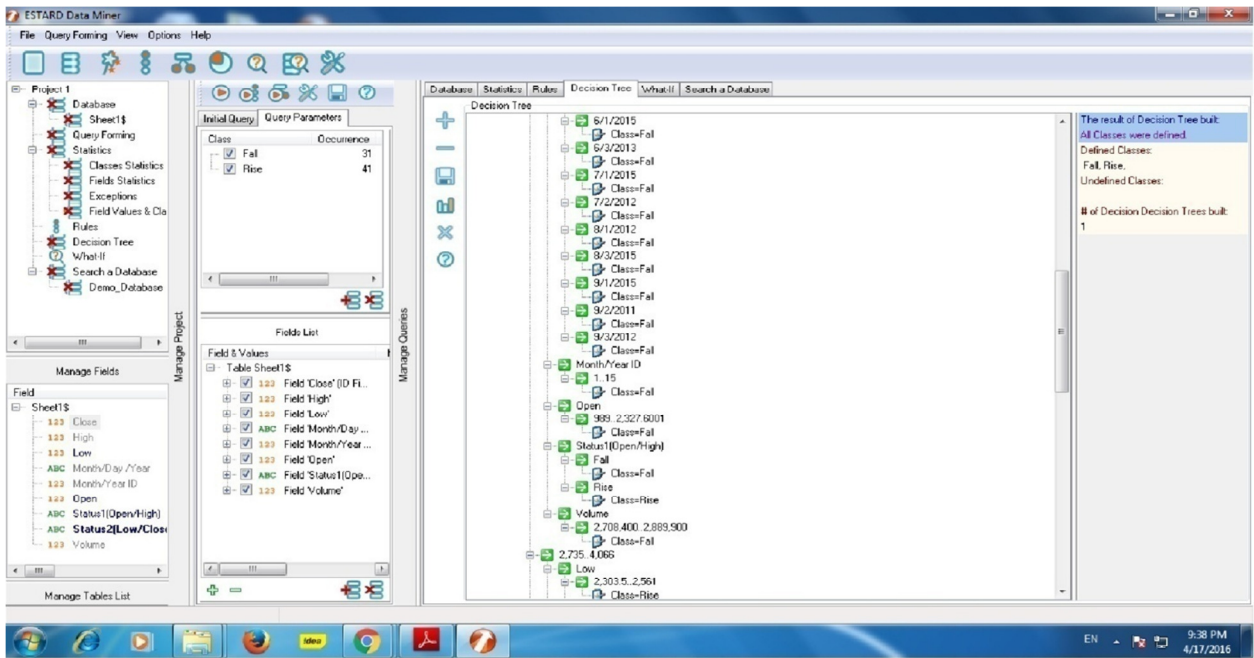Figure 2: Generate decision rule for class attribute status1 (open/high).

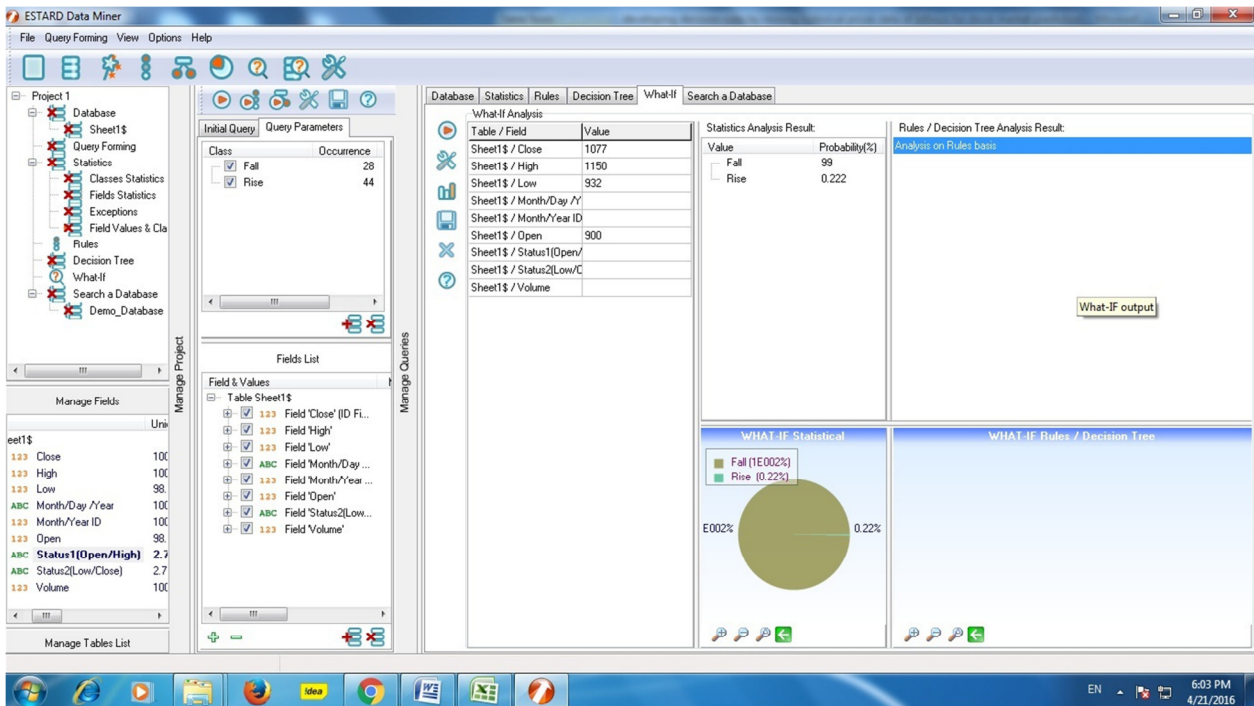Figure 3: Generate decision tree for class attribute status2 (low/close).

Figure 4: Prediction using what-if analyzer

VI.   EXPERIMENT RESULTS

There are number of rules generated from decision tree and Rules classifier of the ESTARD data miner .Also we get many price predictions for their 'Fall' and 'Rise' status according to random selection of attributes from the stock data from What-If analyzer. For Status 1(Open/High) Class value there are 24 rules are generated with 26 for" Fall " and 44 for "Rise " case are met in our experiment .(see figure 2).For Status 2(Close/low) Class value there are 12 rules are generated with 31 for" Fall " and 41 for "Rise " case are met in our experiment. These results are generated during one cycle of experiment. In this way I have applied more cycle of experiment for other attribute also.

The above results show the interestingness patterns in the form of rules. Some of them are as follows:

If the price open with 989- 2490 and get low with 2123-2404, there is 96 % probability that the price classify in the' fall' class.

If price is at high=1110-2735(average=1726) and reach at Low=932-2303 (average=1461) and open with 989-2327(average=1640), then Status1 (Open/High) predicts to' fall' category.

If Price reach at high =2991-4281 then Status1 (Open/High) Predicts price to 'Rise' but in the year 2015 it was classified to 'fall'. Here this and some of other results shows the dependency on particular year measured but the reasons have not been mentioned in our study.

For random selection values if the price open with 2900 and reach at high with 3009 and get low with 2500 and after that get closed with 2800 then there is 99 % probability that the price classify in the' rise 'class. similarly if the price open with 900 and reach at high with 1150 and get low with 932 and after that get closed with 1077 then there is also 99 % probability that the price classify in the' fall 'class.(Fig-4)

VII.   CONCLUSION

In this study, the task is to classify and predict the stock prices of INFOSYS Company that will vary from each day by using the past values of stock price of the company. The rules that were generated from the decision tree and Rule miner can be used   in a system that classify and predict the best action with best timing for the investors, either to buy or sell the stocks on particular day. In this way this study also recommends to use the decision tree classifier that is applied on the historical prices of the stocks to develop decision model that present and conclude the decision about buy or sell the stock. On the basis of analysis of the historical prices of stocks, this type of decision model can be a so useful tool for the investors so that they take the right decision at right time about their stocks in order to extract any useful predictive information from that historical data. We have also seen in our experiment that the results for the proposed model were not accurate or up to the mark because many factors influence the stock market that might be any economic and whether conditions or political events that may affect structure or past sequence pattern tendency of stock market. In this way, there is still huge space for testing and improving the proposed model by assessing the decision model for any company of stock market, for the future work. Also, the other data mining techniques such as association rules and neural networks and clustering [13] can be used to evaluate and investigate the stock market for the classification and prediction. Finally, financial reports and news may be other areas for future work as these can be factors affecting the behavior of the stock markets.

VIII.   REFERENCES

[1] Jan Ivar Larsen, " Predicting Stock Prices Using Technical Analysis and Machine Learning " thesis submitted in Norwegian University of Science and Technology, June 2010.

[2] Bodie−Kane−Marcus, Book-"*equilibrium in capital markets*, Chapter 12*; market efficiency*" The McGraw−Hill Companies, 2001.

[3] Carol Hargreaves, Yi Hao, "Prediction of Stock Performance Using Analytical Techniques" Journal Of Emerging Technologies In Web Intelligence, Vol-5, Issue-2, May 2013.

[4] Jan Ivar Larsen, "Predicting Stock Prices Using Technical Analysis and Machine Learning", Norwegian University of Science and Technology, June 2010.

[5] K. Senthamarai Kannan, P. Sailapathi Sekar, M.Mohamed Sathik and P. Arumugam, "Financial Stock Market Forecast using Data Mining Techniques", Intertnational multiconference of engineers and computer scientists ,Vol-1 Issue-1, pp.17-19, 2010.

[6] Debashish Das and Mohammad Shorif Uddin, "Data Mining And Neural Network Techniques In Stock Market Prediction"- A Methodological Review, International Journal of Artificial Intelligence & Applications (IJAIA),Vol-4,Issue-1, January 2013.

[7] Harish R Pawar, Prasad G Gaikwad, Umesh G Bombale, Dipak D Jagtap and Santosh Durugkar, "Intelligence Stock Forecasting Using Neural Network", International Journal of Computer Sciences and Engineering, Volume-02, Issue-04, pp.103-106, Apr -2014,.

[8] Parth Mody, Advait Marathe, Viral Parekh & Siddhesh Owalekar "An Optimized Approach To Analyze Stock Market" International Journal of Computer Sciences and Engineering, vol-3,Issue-4, pp.19-24,2014.

[9] Ehsan Hajizadeh, Hamed Davari Ardakani and Jamal Shahrabi- "Application of data mining techniques in stock markets" ,A survey; Journal of Economics and International Finance,Vol- 2,Issue-7, pp.109-118, July 2010.

[10] Qasem A. Radaideh, Adel, Eman Alnagi, "Predicting stock prices using data mining techniques" - the international Arab conference on information technology (acit'2013),2013.

[11] Livermore, J. L., "How to trade in stocks, the Livermore formula for combining time element and price" New York, NY: Duel.1940

[12] Stock Market Dataset Available from Yahoo finance India (web-https://in.finance.yahoo.com).

[13] Aditya Joshi, Nidhi Pandey, (Professor) Rashmi Chawla, Pratik Patil-"Use of Data Mining Techniques to Improve the Effectiveness of Sales and Marketing" International Journal of Computer Science and Mobile Computing , Vol-4,Issue-4,pp.81-87,2015.

Author Profile

Dr. Sanjeev Gour pursed Bachelor of Computer Science(Honours) and Master of Science(Elctronics & Comm.) from Devi Ahilya University Indore and Master of Science(Computer Sience) from Barkatullah University Bhopal. He is completed his Ph.D. in Computer Science and currently working as a Professor in Department of Computer Science in Career College,Bhopal. He is a member of Board of Studies & Examination Committee of Computer Science in Bhoapl university also a member of Managing Committee in Computer Cociety of India(Bhopal Chapter). He has published more than 7 research papers in reputed international journals including Thomson Reuters & Scopus (SCI & Web of Science).His main research work focuses on Data Mining. He has 15 years of teaching experience and 4 years of Research Experience.